

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/638,693

1600

DATE: 08/25/2003
TIME: 10:01:19

INPUT SET: S37017.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

RECEIVED

AUG 28 2003

DATA CENTER 1600/2900

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT:

(A) NAME: Innogenetics sa.
(B) STREET: Industriepark Zwijnaarde 7, box 4
(C) CITY: Ghent
(E) COUNTRY: Belgium
(F) POSTAL CODE (ZIP): B-9052
(G) TELEPHONE: 00 32 9 241 07 11
(H) TELEFAX: 00 32 9 241 07 99

(ii) TITLE OF INVENTION: New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.

(iii) NUMBER OF SEQUENCES: 270

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/638,693
(B) FILING DATE: 2000-08-15

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/362,455
(B) FILING DATE: 1995-01-11

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

ENTERED

RAW SEQUENCE LISTING PATENT APPLICATION US/09/638,693

DATE: 08/25/2003

TIME: 10:01:20

INPUT SET: S37017.raw

```

47      (vii) IMMEDIATE SOURCE:
48          (B) CLONE: BR34-4-20
49
50      (ix) FEATURE:
51          (A) NAME/KEY: CDS
52          (B) LOCATION: 1..213
53
54      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
55
56      CTC ACG GAA CGG CTT TAC TGC GGG GGC CCT ATG TTC AAC AGC AAG GGG      48
57      Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
58          1              5              10              15
59
60      GCC CAG TGT GGT TAT CGC CGC TGC CGT GCC AGT GGA GTT CTG CCT ACC      96
61      Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
62          20              25              30
63
64      AGC TTC GGC AAC ACA ATC ACT TGC TAC ATC AAG GCC ACA GCG GCT GCA      144
65      Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
66          35              40              45
67
68      AGG GCC GCA GGC CTC CGG AAC CCG GAC TTT CTT GTC TGC GGA GAT GAT      192
69      Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp
70          50              55              60
71
72      CTG GTC GTG GTG GCT GAG AGT      213
73
74
75
76      Leu Val Val Val Ala Glu Ser
77          65              70
78
79      (2) INFORMATION FOR SEQ ID NO: 2:
80
81          (i) SEQUENCE CHARACTERISTICS:
82              (A) LENGTH: 71 amino acids
83              (B) TYPE: amino acid
84              (D) TOPOLOGY: linear
85
86          (ii) MOLECULE TYPE: protein
87
88          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
89
90      Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
91          1              5              10              15
92
93      Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
94          20              25              30
95
96      Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
97          35              40              45
98
99      Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp

```

RAW SEQUENCE LISTING PATENT APPLICATION US/09/638,693

DATE: 08/25/2003
TIME: 10:01:20

INPUT SET: S37017.raw

```

100          50          55          60
101
102 Leu Val Val Val Ala Glu Ser
103    65          70
104
105 (2) INFORMATION FOR SEQ ID NO: 3:
106
107 (i) SEQUENCE CHARACTERISTICS:
108 (A) LENGTH: 213 base pairs
109 (B) TYPE: nucleic acid
110 (C) STRANDEDNESS: single
111 (D) TOPOLOGY: linear
112
113 (ii) MOLECULE TYPE: cDNA
114
115 (vii) IMMEDIATE SOURCE:
116 (B) CLONE: BR36-23-18
117
118 (ix) FEATURE:
119 (A) NAME/KEY: CDS
120 (B) LOCATION: 1..213
121
122 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
123
124 CTC ACG GAA CGG CTT TAC TGC GGG GGC CCT ATG TTC AAC AGC AAG GGG      48
125 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
126   1          5          10          15
127
128 GCC CAG TGT GGT TAT CGC CGC TGC CGT GCC AGT GGA GTT CTG CCT ACC      96
129 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
130          20          25          30
131
132 AGC TTC GGC AAC ACA ATC ACT TGC TAC ATC AAG GCC ACA GCG GCT GCA      144
133 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
134          35          40          45
135
136 AGG GCC GCA GGC CTC CGG AAC CCG GAC TTT CTT GTC TGC GGA GAT GAT      192
137 Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp
138          50          55          60
139
140 CTG GTC GTG GTG GCT GAG AGT      213
141 Leu Val Val Val Ala Glu Ser
142    65          70
143
144 (2) INFORMATION FOR SEQ ID NO: 4:
145
146 (i) SEQUENCE CHARACTERISTICS:
147 (A) LENGTH: 71 amino acids
148 (B) TYPE: amino acid
149 (D) TOPOLOGY: linear
150
151 (ii) MOLECULE TYPE: protein
152

```

RAW SEQUENCE LISTING PATENT APPLICATION US/09/638,693

DATE: 08/25/2003
TIME: 10:01:20

INPUT SET: S37017.raw

153 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

154
155 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
156   1           5           10           15
157
158 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
159           20           25           30
160
161 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
162           35           40           45
163
164 Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp
165   50           55           60
166
167 Leu Val Val Val Ala Glu Ser
168   65           70
169

```

170 (2) INFORMATION FOR SEQ ID NO: 5:

```

171
172 (i) SEQUENCE CHARACTERISTICS:
173 (A) LENGTH: 213 base pairs
174 (B) TYPE: nucleic acid
175 (C) STRANDEDNESS: single
176 (D) TOPOLOGY: linear
177

```

178 (ii) MOLECULE TYPE: cDNA

180 (iii) HYPOTHETICAL: NO

182 (iii) ANTI-SENSE: NO

184 (vii) IMMEDIATE SOURCE:

185 (B) CLONE: BR36-23-18

187 (ix) FEATURE:

188 (A) NAME/KEY: CDS

189 (B) LOCATION: 1..213

191 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

192
193 CTC ACG GAG CGG CTT TAC TGC GGG GGC CCT ATG TTT AAC AGC AAG GGG      48
194 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
195   1           5           10           15
196
197 GCC CAG TGT GGT TAT CGC CGT TGC CGT GCC AGT GGA GTT CTG CCT ACC      96
198 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
199           20           25           30
200
201 AGC TTC GGC AAC ACA ATC ACT TGT TAC ATC AAA GCC ACA GCG GCC GCA      144
202 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
203           35           40           45
204
205 AAA GCC GCA GGC CTC CGG AGC CCG GAC TTT CTT GTC TGC GGA GAT GAT      192

```

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/638,693DATE: 08/25/2003
TIME: 10:01:20

INPUT SET: S37017.raw

206 Lys Ala Ala Gly Leu Arg Ser Pro Asp Phe Leu Val Cys Gly Asp Asp
207 50 55 60
208
209
210

211 CTG GTC GTG GTG GCT GAG AGT
212 Leu Val Val Val Ala Glu Ser
213 65 70
214
215

213

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 71 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

227 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
228 1 5 10 15
229

230 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
231 20 25 30
232

233 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
234 35 40 45
235

236 Lys Ala Ala Gly Leu Arg Ser Pro Asp Phe Leu Val Cys Gly Asp Asp
237 50 55 60
238

239 Leu Val Val Val Ala Glu Ser
240 65 70
241

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 213 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: BR36-23-20

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/638,693DATE: 08/25/2003
TIME: 10:01:20**INPUT SET: S37017.raw**

Line	Error	Original Text
5	Mandatory Value Not Present	(i) APPLICANT:
6	Unknown or Misplaced Identifier	(A) NAME: Innogenetics sa.
7	Unknown or Misplaced Identifier	(B) STREET: Industriepark Zwijnaarde 7, box 4
8	Unknown or Misplaced Identifier	(C) CITY: Ghent
9	Unknown or Misplaced Identifier	(E) COUNTRY: Belgium
10	Unknown or Misplaced Identifier	(F) POSTAL CODE (ZIP): B-9052
11	Unknown or Misplaced Identifier	(G) TELEPHONE: 00 32 9 241 07 11
12	Unknown or Misplaced Identifier	(H) TELEFAX: 00 32 9 241 07 99

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/09/638,693

DATE: 08/25/2003
TIME: 10:01:20

INPUT SET: S37017.raw

ADDRESSEE
STREET
CITY
STATE
COUNTRY
ZIP
CORRESPONDENCE ADDRESS
CLASSIFICATION

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/09/638,693DATE: 08/25/2003
TIME: 10:01:20

INPUT SET: S37017.raw

Line	Original Text	Corrected Text
4029	(C) INDIVIDUAL ISOLATE: BR33 and BR36	(C) INDIVIDUAL ISOLATE: BR33 and BR36
4053	(C) INDIVIDUAL ISOLATE: HD10	(C) INDIVIDUAL ISOLATE: HD10
4078	(C) INDIVIDUAL ISOLATE: BR36	(C) INDIVIDUAL ISOLATE: BR36
4102	(C) INDIVIDUAL ISOLATE: BR36	(C) INDIVIDUAL ISOLATE: BR36
4129	(C) INDIVIDUAL ISOLATE: HD10	(C) INDIVIDUAL ISOLATE: HD10
4180	(C) INDIVIDUAL ISOLATE: BR36	(C) INDIVIDUAL ISOLATE: BR36
4207	(C) INDIVIDUAL ISOLATE: HD10	(C) INDIVIDUAL ISOLATE: HD10
5297	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5319	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5341	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5362	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5385	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5406	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5428	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5451	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5473	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5495	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5516	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5537	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5559	(C) INDIVIDUAL ISOLATE: GB358 and GB809	(C) INDIVIDUAL ISOLATE: GB358 and GB809
5582	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5604	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549